UNCLASSIFIED

AD NUMBER
AD245989
NEW LIMITATION CHANGE
TO **Approved for public release, distribution unlimited**
FROM **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; OCT 1960. Other requests shall be referred to Office of Naval Research, 800 North Quincy Street, Arlington, VA 22217-5660.**
AUTHORITY
ONR Notice, 8 Jul 1971

THIS PAGE IS UNCLASSIFIED

POST-TRAINING PERFORMANCE CRITERION DEVELOPMENT AND APPLICATION

Generalized Guttman and Thurstone Scales for Electronic Job Performance Evaluation

Prepared For

Personnel and Training Branch

OFFICE OF NAVAL RESEARCH

Under Contract Nonr-2273(00)

Applied Psychological Services

Wayne, Pennsylvania

POST-TRAINING PERFORMANCE CRITERION DEVELOPMENT AND APPLICATION

Generalized Guttman and Thurstone Scales for Electronic Job Performance Evaluation

Prepared For
Personnel and Training Branch
OFFICE OF NAVAL RESEARCH
Under Contract Nonr-2279(00)

Applied Psychological Services
Wayne, Pennsylvania

POST-TRAINING PERFORMANCE

CRITERION DEVELOPMENT AND APPLICATION

Generalized Guttman and Thurstone Scales
for Electronic Job Performance Evaluation

Arthur I. Siegel
Douglas G. Schultz

Prepared for

Personnel and Training Branch

Office of Naval Research

by

APPLIED PSYCHOLOGICAL SERVICES

Wayne, Pennsylvania

under

Contract Nonr-2279(00)

October 1960

ABSTRACT

This is a logical extension of and extrapolation from a series of previous Applied Psychological Services' studies in the development and application of on-the-job criteria for post-training performance evaluation in the Navy. In two of the previous studies, technical proficiency check lists, which meet the Thurstone and Guttman Scalability requirements, were developed for two separate Naval ratings. The purpose of the present study was to develop similar scaled check lists which could be applied inclusively across four ratings which involve electronics work.

A preliminary list of tasks performed by technicians in the four ratings was prepared and utilized in the construction of two check list forms which were analyzed by the Thurstone method of equal-appearing intervals and the Guttman scalogram technique. The relationships between post-training performance proficiency, as measured by the scaled check list, and several other relevant variables, including Naval status, were investigated.

The results seem to support the following conclusions:

1. Skills involved in the Naval ratings of aviation electrician's mate, aviation electronics technician, aviation fire control technician, and TRADEVMAN are scalable by both the Thurstone method of equal-appearing intervals and the Guttman method of scalogram analysis.

2. It is possible to construct a single scaled technical proficiency check list which can be applied to technicians in any of four electronics ratings.

3. As measured by the Scaled Technical Proficiency Check List, TRADEVMEN are significantly less proficient on the electronics tasks common to all four ratings than aviation electrician's mates, aviation electronics technicians, and aviation fire control technicians.

4. As reflected by the scaled lists, the proficiency of electronically oriented technicians on the tasks done in these ratings rises from striker to petty officer second class.

5. The Naval attitudes of electronics technicians, as expressed in a self-report questionnaire, generally are not strongly related to technical fleet proficiency, although attitudes regarding certain aspects of the job may have slight, positive relationships with proficiency.

6. The fleet effectiveness of electronics technicians, as reflected by the scaled lists, is not related to scores on the Navy Basic Test Battery or to technical school grades.

The functional characteristics of two criterion instruments, the Technical Behavior Check List and the Scaled Technical Proficiency Check List, are discussed.

As a result of three Applied Psychological Services' studies, it seems reasonable to conclude that the technical skills involved in Naval ratings are scalable in the same manner as attitudes and the sensory phenomena which have been previously scaled psychophysically and that the scaling can be established either with a single rating or across several related ratings. The scales appear to offer a quick, convenient way of evaluating the post-training technical proficiency of Naval personnel and to provide one basis for judging the effectiveness of technical training programs.

ACKNOWLEDGMENTS

This research was supported by the Office of Naval Research under Contract Nonr-2279(00). We are indebted to Dr. Glenn Bryan and Mr. John Nagay, Personnel and Training Branch, Psychological Sciences Division, ONR, and to Dr. Denzel Smith, Head, Psychological Sciences Division, ONR, for their encouragement and positve comments throughout the program.

Other members of the staff of Applied Psychological Services also participated. Philip Federman and Stephen Benson helped obtain the data and in the development of the research instruments. Terence Connell performed many of the statistical analyses. Gail Gensemer and Estelle Siegel served as project secretaries and prepared the manuscript for publication.

Dr. Douglas Mayo, Staff, Chief of Naval Air Technical Training Command, reviewed a draft version of Chapters I through III of this report and provided advice and assistance at many decision points.

Finally, and most important, various members of the Staff of the Commander Naval Air Force, U. S. Atlantic Fleet performed the necessary in-service liaison for the field work which extended over 24 squadrons and five Naval air stations. Their administrative skill and

effectiveness enabled the easy achievement of what might have been a difficult field program.

Arthur I. Siegel
Douglas G. Schultz

APPLIED PSYCHOLOGICAL SERVICES
October 1960

TABLE OF CONTENTS

TABLE OF TABLES

CHAPTER I

INTRODUCTION

Applied Psychological Services has carried out a series of studies which have had as one important purpose the development of criterion measures for the post-training performance evaluation of enlisted personnel in various Naval aviation ratings. The research has progressed along a line of thinking which first produced Technical Behavior Check Lists (TBCLs) for four ratings (Richlin, Federman, and Siegel, 1958; Siegel, Richlin, and Federman, 1958; Siegel, Richlin, and Federman, 1960; Richlin, Siegel, and Schultz, 1960). The TBCLs were comprehensive, detailed lists of the tasks performed by men in each rating. Psychological scaling techniques were then used in a series of investigations into their applicability to a skill hierarchy. Application of the scaling techniques led to the construction of short, convenient-to-use, post-training performance evaluative instruments for each of two Naval ratings (Siegel and Benson, 1959; Siegel, Schultz, and Benson, 1960). The next logical question to ask was whether a short, scaled fleet performance measurement device could be developed which could be general enough to be used across several ratings. The research described in this report was designed to answer that question.

The Scaling Methods

The scaling approaches employed in the two previous scaling studies (Siegel and Benson, 1959; Siegel *et al.*,1960) were those proposed by Thurstone (1929) and by Guttman (1950). These methods were originally developed primarily for the measurement of psychophysical phenomena and attitudes. One conclusion from the two studies by Siegel and his associates was that skills are scalable in the same manner as attitudes and the sensory phenomena which have been previously scaled psychophysically.

The underlying rationale and application of the techniques to the skill domain were discussed in the two previous reports. Thurstone's requirements for a set of statements or items to be considered a scale were that:

1. the set of statements should all relate to the same psychological object
2. the items should fall along a continuum from "least" to "most"
3. the point at which each statement falls along the continuum should be definable
4. the probability distribution of endorsement of the scale values of the statements should be normal and variance minimal

Guttman considered a set of items to form a scale provided:

1. a person who obtains a higher rank than a second person on a given item is also as high or higher on all other items (reproducibility)

2. the nonfitting responses are well scattered and distributed randomly; clustering would indicate a systematic distortion of the scale pattern (pattern of errors)

3. no or only a few items have almost all responses lumped under a single alternative. For example, an item answered positively by 95 per cent of the respondents could not possess a reproducibility of less than 95 per cent, while an item possessing a 50-50 split could theoretically result in 50 per cent nonfitting responses (range of marginal distribution)

The Thurstone method has been described by Torgerson as typical of the "stimulus-centered or judgment approach" to scaling in which "the systematic variation in the reactions of the subjects to the stimuli is attributed to differences in the stimuli with respect to a designated attribute" (1958, p. 46). The Guttman technique, on the other hand, is taken as an example of the "response approach" in which "variability of reactions to stimuli is ascribed to both variation in the subjects and in the stimuli" (1958, p. 46). Both methods, however, assume an underlying psychological continuum along which stimuli may be ordered.

Need for a Generalized Scaled Check List

Siegel and Benson (1959) demonstrated the scalability, in both the Thurstone and Guttman senses, of the skills involved in the Naval specialty of aviation electronics technician. Siegel *et al* (1960) achieved similar results for the skills involved in the Naval specialty of aviation machinist's mate.

Although the check lists developed in these studies were of value for the post-training evaluation of technicians in a particular rating, it appeared that a short, scaled check list which would apply to several ratings would have wider significance, even greater usefulness, and would also be of considerable interest from the standpoint of scaling theory. Not only would such an instrument allow a more economical means of measuring post-training fleet performance with a minimum of different forms, but the establishment of a common scaled skill hierarchy would also provide a kind of common base across related ratings. This base would have implications for cross-rating evaluations, job task analysis, career planning, the establishment of training requirements across ratings, etc., and possibly might give some basis for grouping across ratings for various purposes. In other words, it might provide a common taxonomy for describing related ratings. Additionally, since there is a constant need for sound, short job oriented instruments for the evaluation of fleet performance in order to provide feedback to training facilities, it seemed to be a worthwhile effort to try to develop the broader-based scaled check lists.

Purposes of the Present Study

The primary purpose of the present study, therefore, was to investigate whether technical proficiency criterion measurement instruments could be constructed which could be applied across several related

Naval ratings (specialties) and which could be scaled across these ratings by both the Thurstone and Guttman techniques. Achieving this purpose embraced two steps: (1) developing behaviorally based items that were general enough to apply to the skills included in the several ratings and yet covered the important duties of each rating, and (2) scaling the items over the several ratings.

Secondary purposes of the study were to establish the relationship of the derived instruments to other pertinent available measures and to determine the predictability of the new instruments from those measures.

Ratings Involved

Electronics was selected as the broad area within which the research would focus. The following five Naval ratings were felt to involve skills of various related types within electronics:

[1] aviation electrician's mate (AE)
[2] aviation electronics technician (AT)
[3] aviation fire control technician (AQ)
[4] aviation guided missileman (GF)
[5] TRADEVMAN (Training Devices Man) (TD)

It was soon determined that there were very few Aviation Guided Missilemen available for study. Accordingly, this rating was dropped and the study was based on the remaining four ratings.

CHAPTER II

DEVELOPMENT OF A GENERAL SCALE FOR ELECTRONIC TECHNICAL SKILLS BY THE THURSTONE EQUAL-APPEARING INTERVAL METHOD

Development of Preliminary Task List

The possibility of contructing a generalized technical skill check list that would scale rested first of all upon evolving an appropriate list of the tasks performed in the several ratings and casting these tasks in a form that would have essentially equivalent meaning for all of the four ratings included. To be useful, the list had to include the tasks which accounted for the major electronic activities of each rating in a form that would result in their being common to the four ratings.

The previous studies of Naval technicians cited in Chapter I gave the investigators considerable insight into the kinds of work done by various kinds of Naval technicians in electronically oriented specialties. The TBCLs which had been developed served as sources of specific suggestions. Consultations were also held with staff members of the Naval Air Technical Training Command. Out of this background a list of 28 tasks was prepared. The form of the items was to present only the basic function in each task, such as "operates" or "calibrates," without reference to any specific equipment. The general directions for the list stated that each item was to be interpreted as a substitution in the sentence "A striker or petty officer third class in the __________ rate ________ equipment which is encompassed by the rate."

The list of 28 tasks was submitted to a total of 28 instructors at the Naval Air Technical Training Command who had squadron experience in either the AE, AT, AQ, or TD ratings. They were asked to indicate in which areas technicians in their rating worked. After the instructors had given their judgments, they were interviewed as to their feelings about the list and for unlisted work areas. In general, the instructors found the lists complete and the terminology acceptable. Although they had a few suggestions to clear up minor confusing points, a majority of these experienced men agreed that 21 of the 28 tasks were done to some degree by strikers or petty officers third class in their rating. Several of the instructors thought that each of the other seven tasks were also worked on but that these seven tasks were performed in the ratings by the higher level personnel. Accordingly, pending further analysis, all 28 tasks were left in the initial, experimental list.

The Preliminary Task List

The 28 tasks which had been developed and checked with experienced instructors were put in a preliminary form. In this form the respondent was presented with an eleven point continuum and asked to indicate where on the continuum each task would fall in difficulty for the average striker. He was told that number *1* on the continuum was to represent the task that is "least difficult for the average striker." Number *7* was to represent "the job that is the "most difficult" for the typical striker. The intermediate numbers were to represent intermediate degrees of difficulty. The respondents were provided with gummed, prenumbered response labels in amounts such that the frequency distribution of the numbers printed on the stickers roughly approximated a normal distribution. The specific frequency distribution of the gummed labels is presented in Table 1. The respondent was asked to place one of the labels next to each listed task.

Table 1

Frequency Distribution of Gummed Response Labels for Preliminary Task List

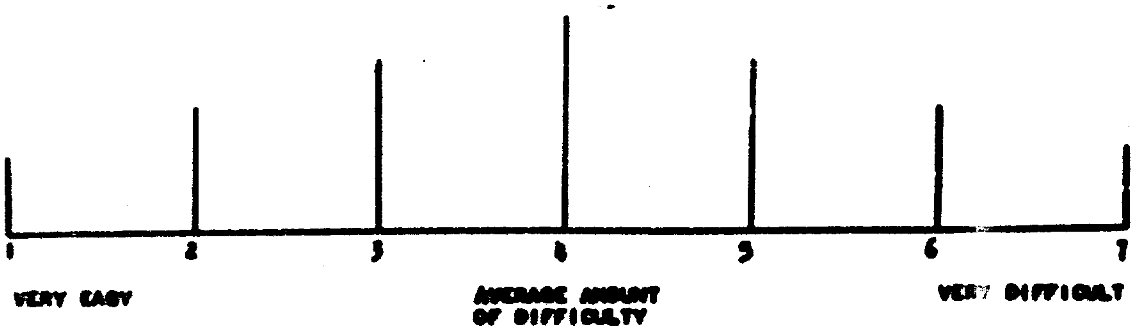
Response Label	Number Provided
1	1
2	3
3	6
4	8
5	6
6	3
7	1

The instructions for the preliminary task list are given in Table 2. The complete form, as used, is presented as Appendix A to this report.

Table 2

Directions for Preliminary Task List

Listed below you will find 28 tasks that are done by strikers in your rating. Read over the 28 tasks and then using the scale below as a guide indicate the degree of difficulty you believe that the typical striker encounters on each of the tasks before he can perform it proficiently. These judgments are to be made in the following manner:



1. Using the stickers provided place the sticker with a 1 on it next to the job that is least difficult for the average striker.

2. Then pick the three jobs that fall in category "two" and put the three stickers with 2's on them next to these items.

3. Now find the job that is the most difficult and place the sticker with the 7 on it next to this item.

4. Find the three items you want to place in category "six" and put these stickers next to these three items.

5. Now with the tasks remaining place them in their proper category by putting the sticker (i.e., 3, 4, or 5) with the category number next to the items.

1	2	2	2	3	3	3	3	3	3
4	4	4	4	4	4	4	4	5	5
5	5	5	5	6	6	6	7		

In previous studies (Siegel and Benson, 1959; Siegel, Schultz and Benson 1960) two other response forms had been used, one asking for the number of checkouts required by the typical striker before he is able to do the task without direct supervision and the other calling for the amount of inservice training required before the typical striker can perform the task proficiently. The analysis of the data in those two studies supported the conclusion that essentially the same scale hierarchy was established by the various question forms. Therefore, in the present study, only the difficulty form was used at this first step.

The 28 items or tasks included in the preliminary task list were:

1. Operating
2. Preflight inspecting
3. Postflight inspecting
4. Inflight inspecting
5. Periodically inspecting
6. Maintaining
7. Removing
8. Repairing
9. Replacing
10. Performing preventative maintenance
11. Trouble shooting/isolating malfunction(s) in
12. Calibrating

13. Aligning

14. Following block diagrams for

15. Using standard schematics for

16. Analyzing standard circuitry in

17. Employing safety precautions on

18. Using proper safety precautions for self when working on

19. Making out failure reports for

20. Using manuals of a technical nature for

21. Using appropriate test equipment for determining malfunction in the

22. Using ASO catalogue for replacement parts for

23. Using mathematical formulas necessary for solving circuit equations for

24. Employing electronic principles involved in maintenance of

25. Knowing relationship of equipment to other related

26. Instructing others in operation of

27. Instructing others in maintenance of

28. Instructing others in the inspection of

Sample

The preliminary task list was administered to 242 enlisted supervisory personnel in the ratings and pay grades described in Table 3 and in the squadrons and locations listed in Table 4.

Table 3

Numbers of Supervisors in Sample by Rating and Pay Grade

	Petty Officer Second Class (2/c)	Petty Officer First Class (1/c)	Chief Petty Officer* (CPO)	Total
Aviation Electrician's Mate (AE)	26	21	21	68
Aviation Electronics Technician (AT)	28	24	28	80
Aviation Fire Control Technician (AQ)	3	22	19	44
TRADEVMAN (TD)	14	19	17	50
TOTAL	71	86	85	242

* Included in the CPO group were 17 Master Chiefs and 2 Senior Chiefs

Table 4

Numbers of Supervisors in Sample by Location and Squadron

Location	Squadron	Number
Cecil Field	FAETULANT	1
	FASRON 9	22
	FITRON 14	2
	VA 46	7
	VF 11	10
	VF 14	7
	VF 174	17
	VFP 62	11
Jacksonville	AEWRON 4	1
	VA 44	29
	VAP 62	7
	VP 16	14
	VW 4	21
Norfolk	FAETULANT	25
	FASRON 3	4
	FASRON 102	3
	VR 22	5
	VRF 31	2
Quonset Point	FAETULANT DET. 3	14
Sanford	FASRON 51	3
	VAH 3	14
	VAH 5	8
	VAH 7	13
	VAH 11	2
		242

Administration

The preliminary task list was administered to groups of supervisors at each base. A full explanation of the purpose of the study and instructions on how to use the gummed labels were given to the raters before the preliminary task list was distributed. The supervisors were asked to complete a Sailor's Naval Attitude (SNA) Inventory after they completed the preliminary task list.

Each group session took approximately one hour.

Results

Using the response data obtained from the administration of the preliminary task list to the 242 supervisors, the median and interquartile range were calculated for each item (or task). These provided the scale (S) and deviation (Q) values needed for establishing a scale according to Thurstone's method of equal-appearing intervals. The results are plotted in Figure 1. In examining this figure, it should be remembered that the rater was forced to respond on a seven point scale and to normalize approximately the distribution of his responses. The lowest scale value obtained was 1.52 for item 7 ("removing") and the highest was 6.06 for item 11 ("trouble shooting/isolating malfunction(s) in"). While this is a good range of S values, the very extreme positions are not represented, a finding consistent with the previous two studies. The Q values are fairly constant over the entire range of S values, with perhaps a slight suggestion that they are higher for the more difficult tasks.

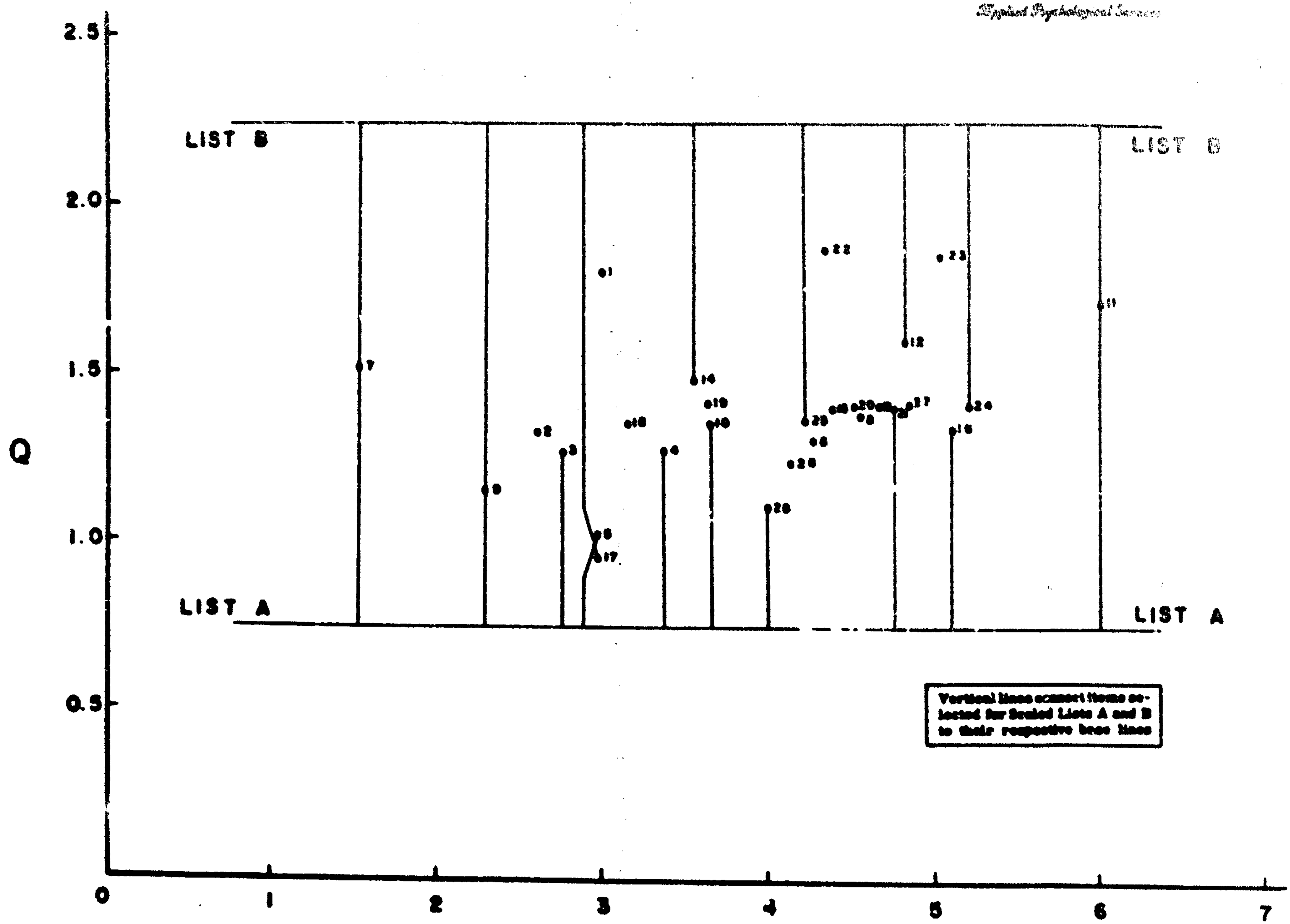


FIGURE I S AND Q VALUES FOR ITEMS IN THE PRELIMINARY TASK LIST

Item Selection for a Thurstone Type Scale

In order to select a subset of items (tasks) which would form a Thurstone equal-appearing interval scale, items were sought which would:

1. represent all values along the psychological "difficulty" continuum (scattering of S values)
2. have minimum Q values
3. sample all technical areas performed in the ratings involved

Since it was hoped that two parallel scales could be constructed, two sets of items were selected. Because of the single items available at the extremes of the S value distribution, it was necessary to accept three items ("removing," "replacing," and "trouble shooting/isolating malfunction(s) in") for both sets, thus introducing common elements into any scores based on these scales.

The selected tasks for the two scales, with their S and Q values, are presented in Table 5. The groups of selected tasks are also indicated in Figure 1. These data confirm the conclusions of the previous two studies that skills are scalable in the same manner as attitudes and the sensory phenomena which have been previously scaled psychophysically. Here, moreover, a single scale was derived which was applicable to a group of four different but related Naval electronics ratings, rather than to just a single rating.

Table 5

S and Q Values for Selected Items

	Scaled List A S	Scaled List A Q	Scaled List B S	Scaled List B Q
Removing (7)*	1.52	1.51	1.52	1.51
Replacing (9)	2.31	1.16	2.31	1.16
Postflight inspecting (3)	2.77	1.27		
Employing safety precautions on (17)			3.00	0.96
Periodically inspecting (5)	3.00	1.02		
Inflight inspecting (4)	3.41	1.28		
Following block diagrams for (14)			3.59	1.48
Performing preventative maintenance (10)	3.70	1.35		
Instructing others in the inspection of (28)	4.03	1.10		
Knowing relationship of equipment to other related (25)			4.27	1.36
Using appropriate test equipment for determining malfunctions in the (21)	4.79	1.41		
Calibrating (12)			4.93	1.00
Analyzing standard circuitry in (16)	5.15	1.34		
Employing electronic principles involved in maintenance of (24)			5.24	1.40
Trouble shooting/isolating malfunction(s) in (11)	6.05	1.72	6.05	1.72

* Number in parenthesis is the item number in the preliminary task list.

CHAPTER III

GUTTMAN TYPE SCALE ANALYSIS OF TECHNICAL SKILLS CHECK LISTS

The Guttman method for scaling items has been discussed in detail in the two previous Applied Psychological Services' reports of research in skill scaling (Siegel and Benson, 1959; Siegel *et al.*, 1960). The second of these summarizes Guttman's basic approach as follows:

> "If a single psychological variable underlies the responses to a set of items, then it should be possible to order individuals on this variable by the hierarchy of responses to the items. Conversely, to the extent that a hierarchy is present, to that extent may an underlying (latent) single dimension be assumed" (Siegel *et al.*, 1960, p. 21).

In using the technique, one seeks an order or hierarchy in the given set of items such that each individual's overall rank is directly related to the highest item in the set which he endorses or passes. If such an order can be found, the set of items is said to scale and to measure a single latent dimension.

Evaluation Form

In order to establish scalability in the Guttman sense, the 28 tasks included in the previously described preliminary task list were put in a form which would allow for evaluations of individuals rather than tasks.

The directions for this form differed from those of the preliminary task list in two respects: (1) they were oriented in terms of a *specific man* whom the rater had supervised rather than the *average striker*, and (2) they asked whether the man being rated is *checked out* on the task rather than how *difficult* the task is for the typical striker.

The response alternatives available to the rater for each task for each man evaluated were:

1. Has worked on task and is checked out
2. Has worked on task and is not checked out
3. Has not worked on task

The full instructions are presented in Table 6. The complete form constitutes Appendix B of this report.

Table 6

Directions for Evaluation Form

For each of the tasks listed indicate (by placing a check mark in the appropriate column) whether or not the man you are evaluating has been checked out as being proficient (i. e., is he capable of doing the task "on his own" without direct supervision?) If the man you are rating has not been checked out as being proficient on a task because he has not worked on this task, this should be indicated by a check mark in the third column.

Examples:

1. If the man has been checked out on a task place a check mark (✓) in column one (1).
2. If he has not been checked out as being proficient but has actually worked on this task place a check mark (✓) in column two (2).
3. If he has not been checked out as being proficient because he has not performed this task, place a check mark (✓) in column three (3) next to the appropriate item.

Be sure that you give an answer for every task that is listed. You may not be certain about some items, but answer every item on the basis of what appears to be the most nearly correct answer for the man you are rating.

Sample

A total of 181 technicians were evaluated through this form. The technicians were distributed in the ratings and pay grades shown in Table 7 and in the squadrons and locations shown in Table 8.

Table 7

Numbers of Technicians in Sample by Rating and Pay Grade

	Striker (AN)	Petty Officer Third Class (3/c)	Petty Officer Second Class (2/c)	Petty Officer First Class (1/c)	Total
Aviation Electrician's Mate (AE)	10	38	-	-	48
Aviation Electronics Technician (AT)	16	35	-	-	51
Aviation Fire Control Technician (AQ)	10	22	7	2	41
TRADEVMAN (TD)	7	23	10	1	41
TOTAL	43	118	17	3	181

Table 8

Numbers of Technicians in Sample by Location and Squadron

Location	Squadron	Number
Cecil Field	FAETULANT	1
	FASRON 9	11
	FITRON 14	1
	VA 46	3
	VF 11	6
	VF 14	16
	VF 174	12
	VFP 62	4
Jacksonville	AEWRON 4	2
	VA 44	21
	VAP 62	7
	VP 18	16
	VW 4	8
Norfolk	FAETULANT	5
	FASRON 3	3
	FASRON 102	2
	VR 22	5
	VRF 31	1
Quonset Point	FAETULANT DET. 3	28
Sanford	FASRON 51	2
	VAH 3	13
	VAH 5	9
	VAH 7	3
		181

Administration

This form was administered to the same supervisors in groups as the preliminary task list. The supervisory personnel were described in Tables 3 and 4. Each rater was asked to evaluate a technician he had supervised; the technician was not necessarily the best or the poorest man he had had under him.

Analytic Method

As in the two earlier scaling studies, the method of scalogram analysis proposed by Green (1956) was employed. This analytic method, an extension of Guttman's technique, places emphasis on a single statistic, the index of consistency, I, in place of the several requirements for scalability proposed by Guttman. I relates the obtained reproducibility (which Green computes from summary statistics) to that expected by chance. He suggests that I should be .50 or greater, if the set of items is to be considered a scale in the Guttman sense. Green writes:

> "This criterion appears to give roughly comparable results to the many criteria used heretofore and will be helpful to those who desire to create a dichotomy of scales *vs*. nonscales" (1956, p. 87).

It should be recognized, however, that Green's selection of a specific value of I for the break between scales and nonscales is an arbitrary matter. In general, the higher the I, the greater the confidence that can be placed in the scalability of the item set.

One problem was whether to attempt to scale the items for the technicians in each of the four ratings separately or to work with the data from all four groups combined.

Establishing the scalability of an item set over the four ratings separately would not thereby establish its scalability over the total group. Guttman writes:

> "A universe may not form a scale for the total population, but still form a scale for subgroups of that population" (1950, p. 83).

This is termed the "relativity" of scales and Suchman argues that:

> "Scales are relative both to time and to populations" (1950, p. 168).

However, it did seem reasonable to begin with an analysis of the entire sample since a finding of scalability at that level would lead to the conclusion that the item set also scaled within each rating group. On this point Guttman states:

> "...if a scale is obtained for a cross section of the population, then that same scale pattern necessarily holds for all major subgroups" (1950, p. 83).

Therefore, in the present study, the analysis first treated the response data from all four ratings taken together, with the thought that if scalability was not established at that level, the analysis would then proceed to various combinations of three or two ratings.

The items or tasks to be tested for Guttman scalability were those included in Scaled List A (10 items) and Scaled List B (3 items) which had been selected as described in Chapter II because they formed Thurstone equal-appearing interval scales. Guttman has not provided any method for the preliminary selection and ordering of a set of items. In this study, as in the previous two, this was accomplished by using the Thurstone analysis as a first step in the Guttman analysis.

Since Green's method requires dichotomous scoring, the "not checked out" and "not worked on" categories of the evaluation form were considered equivalent, as opposed to the "checked out" response.

Results

When the responses to the 10 items of the Thurstone Scaled List A were subjected to a Guttman analysis, a reproducibility figure of .903 and an I of .424 were obtained. Since the I value did not reach Green's critical level of .50, the next step was to consider dropping from the analysis the technicians of one of the ratings. Review of the data suggested that the responses of the TRADEVMAN differed from the responses of the other three ratings more than those of the three did from one another. But the I obtained without the TRADEVMEN, although higher, was only .469.

Since it seemed apparent that the disturbing influence was in the items rather than in the sample, the next step was to drop some anomalous items. There is some disagreement about the wisdom of this procedure. Green, for example, writes:

"If a set of items does not scale, the possibility exists of rejecting one or two poor items, and then achieving a scale. Guttman is chary of this procedure, preferring to say that the universe is not scalable. However, it seems possible to have perfectly good items with the wrong form for the Guttman scale. To this author, the possibility of rejecting items seems to be a necessary part of any method of (attitude)* measurement (1954, p. 357).

Torgerson takes the same point of view as Green. He says:

"...further way to increase 'scalability' is simply to discard or revise the offending items. While Guttman in general frowns on this procedure it would seem to be necessary in many cases" (1958, p. 330).

Tasks 3 ("postflight inspecting"), 4 ("inflight inspecting"), and 28 ("instructing others in the inspection of") were eliminated from Scaled List A and the remaining seven items analyzed. The results are shown in Table 9. The obtained I of .580 is high enough to conclude that the seven items involved form a scale in the Guttman sense. These seven items are henceforth referred to as the Scaled Technical Proficiency Check List, Form A (STPCL A).

* parentheses ours

Table 9

Results of Scalability Analysis of STPCL A

Reproducibility (Rep.)	.942
Reproducibility expected by chance (Rep. I)	.862
Index of consistency (I)	.580

The results from the Guttman analysis of the eight items in the Thurstone Scaled List B are presented in Table 10. The I of .549 indicates that these items also constitute a Guttman scale. These eight items are henceforth referred to as the Scaled Technical Proficiency Check List,Form B (STPCL B).

Table 10

Results of Scalability Analysis of STPCL B

Reproducibility (Rep.)	.936
Reproducibility expected by chance (Rep. I)	.858
Index of consistency (I)	.549

The value of I for STPCL A is probably inflated to some extent by the fact that, in deciding upon which items to drop, the response matrix was examined and, therefore, some advantage was taken of chance relationships in the data. Although I for STPCL A should be checked in another population sample to determine its value more accurately, the fact that STPCL B, from which no items were eliminated, scaled would suggest that the I given above for List A is not a gross overestimation.

Discussion

The fact that it was possible in the present study to establish scales over four related but different Naval ratings has several significant implications. It is apparently possible to generalize a function by divorcing it from specific equipment and still retain its meaningfulness in different situations. This was the central problem faced in writing the items or task descriptions. That it is possible to scale these items means that the technicians in the several ratings involved are all doing the general, basic tasks in the same order (as judged by the ranking in terms of task proficiency) and that if they are checked out on one task on a scaled list, it can be assumed that they are proficient on the tasks which are ranked below that one on the list. This would seem to be of value in understanding the basic structure of the four ratings and the interrelationships among them and to have significance for the development of training programs. Such a hierarchy also seems to be of value for describing the work performed by the men in the ratings, the sequence of technical development, and for rating structuralization.

The tasks in the STPCL, Forms A and B, may be arranged in three groups in terms of the proportion of technicians who were proficient on them. On none of these tasks were less than 25% of the men checked out. Therefore, assignment to a group may be said to depend on whether 75 to 100%, 50 to 75%, or 25 to 50% of the technicians were described as being checked out. Tables 11 and 12 show the percentage of strikers

Table 11

Items in STPCL A by Percentage of Technicians Who Could Perform Task Without Direct Supervision

75-100%	50-75%	25-50%
(7)*Removing (9) Replacing (5) Periodically inspecting		
	(10) Performing preventative maintenance (21) Using appropriate test equipment for determining malfunctions in the (16) Analyzing standard circuitry in	
		(11) Trouble shooting/isolating malfunction(s) in

* Number in parentheses is the item number in the preliminary task list and evaluation form.

Table 12

Items in STPCL B by Percentage of Technicians Who Could Perform Task Without Direct Supervision

75-100%	50-75%	25-50%
(7)*Removing (9) Replacing (17) Employing safety precautions on		
	(14) Following block diagrams for (25) Knowing relationship of equipment to other related	
		(12) Calibrating (24) Employing electronic principles involved in maintenance of (11) Trouble shooting/isolating malfunction(s) in

* Number in parentheses is the item number in the preliminary task list and evaluation form.

that could perform each of the tasks in Scaled Technical Proficiency Check Lists A and B. In view of the findings of the earlier studies that judgments regarding inservice training required and task difficulty were found to rank in the same order as judgments regarding proficiency, it would seem reasonable to say that the tasks in the 75-100% group are those which are easiest and require the least inservice training, while those in the 25-50% group are those which are most difficult and require the most inservice training.

In several earlier studies (Richlin *et al.*, 1958; Siegel *et al.*, 1958; Siegel *et al.*,1960; Richlin *et al.*, 1960) the criterion instrument was a Technical Behavior Check List (TBCL). This was a comprehensive list of all the tasks performed in a rating. The TBCL score was an evaluation of how well, on the average, the technician was performing those tasks on which he had had an opportunity to work. If he had done only a few tasks but performed them well, he obtained a high score. If he had done many tasks but all of them less well, he obtained a lower score. In evaluating an individual TBCL score, the scoring method employed must be considered. For example, a man of limited ability may be restricted to working on a few rather simple tasks which he quickly learns to do. On the other hand, a very competent man may be called upon to work on many tasks at all levels of difficulty, some of which may require considerable practice before a man can accomplish them "on his own." Provided this aspect is kept in mind, the TBCL is effective in assaying how well a technician has done his work.

The scaled technical proficiency check lists, on the other hand, evaluate the overall status of the technician with reference to all the tasks normally performed by men of equivalent pay grade and rating. These lists contain only a relatively small number of items so that they can be quickly and easily completed. But the scales are so constructed that the score obtained from them can be generalized in meaning to the "universe" of tasks of which they are representative. Thus, the STPCL, Forms A and B, identify where the technician stands relative to typical developmental progress within his rating and in related ratings.

The TBCL should prove to be valuable for providing molecular information on how well technical school graduates are performing as a group and thus aid in the evaluation of the effectiveness of training programs. Each task in the comprehensive list can be checked from the standpoint of the proficiency of graduates on that specific aspect of the job. When this information is tied in with the training, it should provide a sound basis for recommendations regarding changes in the emphasis placed on various parts of the training program.

The STPCL also has value for training evaluation purposes. The data from the selected items which constitute the STPCL may be examined to determine what proportion of the group is checked out on these particular tasks; changes in training may then be recommended where greater or less emphasis is desired in the fleet Also, when an experienced technician in one rating is being retrained in another related specialty, his score on the STPCL which has been scaled over several ratings should suggest the approximate level at which the retraining should commence. The validity of this use of the STPCL depends upon the generalizability and transferability of the skills involved.

The two types of instruments, therefore, measure somewhat different aspects of job performance and serve different purposes. In the study by Richlin *et al.*, (1960), it was found that the correlation between the two kinds of measures, as developed for aviation electronics technicians, was positive but low; yet both types of scores had roughly equivalent correlations with "A" school grades.

One advantage of the scaled lists is that they are simple and convenient to use. Also, the availability of two scaled lists may be helpful in certain situations such as reevaluations within a short period of time, although it must always be borne in mind that there is some identity of content between the two lists. They may, however, be considered as roughly equivalent or parallel forms.

CHAPTER IV

INTERRELATIONSHIPS AND PREDICTION OF TECHNICAL FLEET PERFORMANCE

Intercorrelations

When any new measurement instrument is developed, it is of interest to learn how its scores relate to other known measures. Such information is useful in judging the theoretical significance and practical value of the new instrument. In order to determine some relationships of this kind for the two Scaled Technical Proficiency Check List forms, scores on several other relevant variables were obtained for the technicians who were evaluated by their supervisors.

The final class average (FCA) attained in "A" school was procured for each man. Scores from the following Navy Basic Test Battery aptitude tests were also available:

General Classification Test (GCT)
Arithmetic Test (ARI)
Mechanical Test (MECH)
Clerical Test (CLER)

In addition, a Sailor's Naval Attitude (SNA) Inventory was administered to each technician who was evaluated. The SNA Inventory is a self-report inventory which measures the sailor's perception of certain job aspects purported to be important to industrial employee motivation. Its development, content, and characteristics have been fully described in another report (Siegel and Schultz, 1960). In addition to a total score, the inventory yields part scores in each of the following five areas:

Intrinsic Job Satisfaction (I): including those job satisfaction aspects which are derived from direct performance of the work itself and which would tend to be constant for the job no matter where the work is performed

Supervisory Relationships (S): including the relations that exist between the worker and his immediate supervisor

Social Aspects of the Job (SA): including those job aspects involved in the relationships of a sailor with other sailors, especially those of comparable rating

Opportunity for Advancement (O): including those job aspects which the individual sees as potential sources of bettering his economic position, organizational status, or professional experience

Extrinsic Aspects of the Job (X): This score includes:

(a) working conditions: the sailor's reactions to the physical aspects of the working environment which are not necessarily part of the work

(b) benefits: the sailor's reactions to those phases of the service which attempt to protect him against such emergencies as illness or old age

(c) wages: the sailor's reactions to his pay

(d) security: the sailor's reactions to those features of the service which lead to continued employment and/or to vocational opportunity on leaving the service.

The intercorrelations among all these variables, along with means and standard deviations, are presented in Table 13. The N for Table 13 was reduced to 104 because of cases for whom only a part of the data was available. This table is identical to Table 19 on page 38 of the report by Siegel and Schultz (1960). It is presented here again because the two Scaled Technical Proficiency Check List forms were only an incidental part of that report, and it seems worthwhile to examine the table here from the standpoint of the correlations involving those forms, i.e., the correlations in the last two rows of Table 13. Comments on the remainder of the table will not be repeated.

The first six coefficients in each of the last two rows of Table 13 reflect the relationship between the several SNA Inventory scores and the two scaled lists. Out of twelve coefficients, seven are statistically significant at the 5% level of confidence, although all seven are only between .20 and .30. The SNA Inventory scores that produce these significant correlations are Intrinsic Job Satisfaction, Supervisory Relationships, Extrinsic Aspects of the Job, and total score on all the parts.

The correlation coefficients between the two scaled lists and the four "basic battery" test scores vary around and are close to zero. The correlations between FCA and the lists are low and not statistically significant.

Table 13

Means, Standard Deviations, and Intercorrelations Among SNA Inventory Scores, Basic Battery Test Scores, School Grades, and Scaled Check List Scores for ATs, AQs, AEs, and TDs Combined (N = 104)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. I	-												
2. S	.64*	-											
3. SA	.70*	.65*	-										
4. O	.46*	.30*	.43*	-									
5. X	.78*	.55*	.59*	.45*	-								
6. Total Inventory	.90*	.83*	.85*	.57*	.82*	-							
7. GCT	-.10	.02	-.07	-.06	-.07	-.06	-						
8. ARI	-.10	-.05	-.10	.09	-.08	-.08	.38*	-					
9. MECH	.06	.03	.02	.00	.05	.04	.30*	.19	-				
10. CLER	.02	-.09	.03	-.03	-.06	-.03	.21*	.41*	.15	-			
11. FCA	.05	.07	.03	-.06	-.02	.04	.32*	.34*	.26*	.20*	-		
12. STPCL-A	.24*	.25*	.12	.07	.13	.22*	-.03	.12	.02	-.10	.18	-	
13. STPCL-B	.25*	.29*	.16	.14	.21*	.27*	.01	.15	.05	-.07	.15	.89*	-
Mean	12.73	10.38	8.65	7.03	5.88	44.66	60.29	56.31	56.67	50.51	75.38	4.88	5.47
S.D.	3.05	3.56	2.80	1.40	2.20	10.82	6.59	5.01	7.07	8.25	6.07	2.04	2.13

* Significant at 5% level

The two scaled lists have a high intercorrelation, which reflects to some extent the effect of the three tasks common to the two lists. As an approximate and somewhat inflated estimate of the "parallel form" reliability of the STPCL forms, this value of .89 suggests that they may possess satisfactory reliability.

In summary, it appears that job proficiency as measured by the Scaled Technical Proficiency Check List may be slightly related to some job attitudes but that proficiency is not related to aptitude test scores or, to any significant degree, to "A" school grades. This finding for the Navy parallels that of Thorndike and Hagen (1959) who found aptitude tests to be useful for selecting occupational memberships but not for predicting success *within* an occupational classification.

Prediction of Technical Fleet Performance

In order to determine how well on-the-job technical proficiency could be predicted from the variables investigated in this study, several multiple correlation coefficients were computed from the regressions of STPCL, Form B, scores on various combinations of predictors. The simplest and most efficient group of predictive variables consisted of the Intrinsic Job Satisfaction (I) and Supervisory Relationships (S) scores of the Sailor's Naval Attitude (SNA) Inventory and the Arithmetic Test (ARI) of the Navy Basic Test Battery. The multiple regression equation in raw score form which resulted from this combination was:

$$Y = .091X_1 + .129X_2 + .074X_3 - 1.340$$

where Y = STPCL, Form B, score

X_1 = Intrinsic Job Satisfaction (I) score

X_2 = Supervisory Relationships (S) score

X_3 = Arithmetic Test (ARI) score

Use of the weights given in the above equation resulted in a multiple correlation of .35 between STPCL B and the three predictors. This indicates that STPCL B scores can be predicted from these variables with only a moderate degree of accuracy. In fact, the Intrinsic Job Satisfaction score alone had a correlation of .29 with the scaled list scores; thus, the addition of the other two variables increased the correlation only slightly.

Obviously, none of the factors included in this study predicted job proficiency, as measured by the two scaled check list forms, with any high degree of effectiveness. If the check lists are accepted as instruments measuring significant job behavior, new predictors should be uncovered in the future. It may be, for example, that particular sections or certain aspects of the pre-fleet technical training have a direct bearing on fleet effectiveness, even though the overall school grades do not. On the other hand, attitudes and drives may play a greater part on the job than in school grades (or the attitudes and drives which are important for success in school may be different from those which are important for success on the job). This possibility is suggested by the fact that the SNA Inventory scores had higher correlations (Table 13) with the Scaled Technical Proficiency Check List than with "A" school grades. If this is true, perhaps other personal characteristics could be identified which would be highly related to job proficiency. In any case, additional variables should be investigated for purposes of prediction of on-the-job technical proficiency in the electronics ratings studied.

Comparisons Among Ratings and Pay Grades

As a check on the reasonableness of the STPCL scores, means and standard deviations were computed separately for the various ratings and pay grades represented among the men evaluated by their supervisors in this study. Analyses of variance were also carried out among ratings and among pay grades on each of the STPCL forms. The results are presented in Tables 14-17.

Table 14

Means, Standard Deviations, and Analysis of Variance Results for Four Ratings on Scaled Technical Proficiency Check List, Form A, Scores

Rating*	N	Mean	Standard Deviation
AE	48	5.02	1.87
AQ	32	4.91	1.61
AT	51	5.00	1.92
TD	30	3.27	2.35

* Strikers and Petty Officers Third Class only

Source	df	ss	ms
1. Between ratings	3	72.32	24.11
2. Within ratings	157	603.57	3.84
3. Total	160	675.89	-

$F = 6.28$ $P < .01$

Table 15

Means, Standard Deviations, and Analysis of Variance Results for Four Ratings on Scaled Technical Proficiency Check List, Form B, Scores

Rating*	N	Mean	Standard Deviation
AE	48	5.17	1.75
AQ	32	5.41	1.66
AT	51	5.92	2.10
TD	30	3.63	2.37

* Strikers and Petty Officers Third Class only

Source	df	ss	ms
1. Between ratings	3	101.42	33.81
2. Within ratings	157	629.05	4.01
3. Total	160	730.47	-

$F = 8.43$ $\qquad$ $P < .01$

Table 16

Means, Standard Deviations, and Analysis of Variance Results for Three Pay Grades on Scaled Technical Proficiency Check List, Form A, Scores

Pay Grade	N	Mean	Standard Deviation
2/c	17	6.00	1.49
3/c	118	4.90	2.08
AN (Striker)	43	[illegible].02	1.81

Source	df	ss	ms
1. Between pay grades	2	51.55	25.78
2. Within pay grades	175	689.76	3.94
3. Total	177	741.31	-

$F = 6.54$ $P < .01$

Table 17

Means, Standard Deviations, and Analysis of Variance Results for Three Pay Grades on Scaled Technical Proficiency Check List, Form B, Scores

Pay Grade	N	Mean	Standard Deviation
2/c	17	6.82	1.50
3/c	118	5.43	2.15
AN (Striker)	43	4.44	1.88

Source	df	ss	ms
1. Between pay grades	2	73.07	36.54
2. Within pay grades	175	738.03	4.22
3. Total	177	811.10	-

$F = 8.66$ $P < .01$

The analysis of variance findings included in Tables 14 and 15 indicate that the means among the strikers and petty officers third class in the four ratings vary more than would be expected by chance alone. The TDs are also more variable in job proficiency *within* their group than are the ATs, AEs, and AQs. Examination of the tables reveals that the AT, AE, and AQ means are all very similar, but that the TD mean is about a standard deviation lower than the other three. This supports the conclusion that the average TD is less proficient than the average sailor in the other three ratings in the basic tasks common to the four ratings. This conclusion is particularly interesting in view of earlier findings (Siegel and Schultz, 1960) that, on the average, TRADEVMEN have more "favorable" job attitudes than ATs, AEs, and AQs; although the job seems to be more satisfying for the TDs, they appear to be less capable in the basic skills called for in the rating. There are a number of possible explanations for this situation. For example, TDs may be called upon to do a greater variety of tasks, many of which are not common to the electronics group and, therefore, even though their overall proficiency is high, they are not as proficient as the other three ratings on the electronics tasks as such. Or, it is possible that electronics work on training devices is basically different in certain respects from similar work on aircraft. Whatever the explanation, the findings based on STPCL scores appear to be quite definite.

Tables 16 and 17 show that the means of the men in the pay grades of striker, petty officer third class, and petty officer second class, also vary more than would be expected by chance. Since a sailor's pay grade reflects his level of aptitude, training, experience, knowledges, skills and responsibilities, his ability to do the basic tasks of his rating could be expected generally to increase as he rises in pay grade. This rise is reflected in the data of Tables 16 and 17; the increase in average STPCL scores from striker to petty officer third class is approximately half a standard deviation and is about the same from petty officer third class to petty officer second class.

The results of the comparisons among ratings and pay grades appear to be reasonable and so lend some support to the soundness of the STPCL as a criterion instrument.

CHAPTER V

SUMMARY AND CONCLUSIONS

Summary

The research described in this report is one part of a series of studies directed toward the development of on-the-job criteria for the post-training performance evaluation of enlisted personnel in several Naval aviation ratings. For the purpose of producing criteria which would reflect the developmental stage or level of the Naval technician, Siegel and Benson (1959) and Siegel *et al.*, (1960) scaled, by two techniques, the skills involved in two separate Naval aviation ratings. The purpose of the present study was to develop similar scaled task check lists which could be applied across several related ratings. It was felt that criterion instruments of this kind would be easy and economical to use, would demonstrate a way of keeping the number of different forms to a minimum, and would have implications for cross-rating evaluations and actions.

The following four Naval ratings were selected for study because they involved skills in the broad area of electronics:

aviation electrician's mate (AE)
aviation electronics technician (AT)
aviation fire control technician (AQ)
TRADEVMAN (TD)

As in the two previous scaled check list studies, the scaling methods described by Thurstone (1929) and Guttman (1950) were used. First, a comprehensive list of the tasks performed in the four ratings was developed and submitted for criticism and suggestions to a group of instructors who had squadron experience in the ratings. The revised list was then administered to enlisted supervisory personnel, with instructions to estimate the degree of difficulty of each task for the average striker in the rating. These responses formed the basis for the analysis by the Thurstone method. Utilizing the results of this first phase of the study as groundwork, the supervisory evaluations of a group of technicians, in terms of whether the man was checked out on each task, were then analyzed by the Guttman technique.

The relationships between post-training performance proficiency and several other relevant variables, including Naval status, were also investigated. In addition, prediction of job proficiency was studied by a multiple regression analysis.

For the scaling investigations, men from all four ratings were treated as one composite. Two roughly parallel lists of ten and eight selected tasks were found to scale from the standpoint of the Thurstone standards. The eight-item check list also scaled according to Green's (1956) criterion for the establishment of a Guttman type of scale. The other list met Green's criterion only after three tasks were deleted.

The two Guttman scaled check lists were found to have a high interrelationship due partly to the presence of three items, at the extremes of the continuum, common to both lists.

Several scores from a self-report job attitude inventory were found to have slight positive correlations with job proficiency as measured by the two forms of the Scaled Technical Proficiency Check List (STPCL). However, no statistically significant correlations were found between the check lists and either scores from the Navy Basic Test Battery or "A" school grades. A multiple regression equation was derived which produced a weighted combination of two attitudinal subtest scores and one aptitude test. The weighted sum had a correlation of .35 with STPCL, Form B.

Comparisons on STPCL scores among the four ratings represented in the subjects revealed that the TRADEVMEN attained mean scores significantly lower than the AEs, AQs, and ATs. This contrasted with an earlier finding (Siegel and Schultz, 1960) that TRADEVMEN had generally more favorable job attitudes. There was a regular increase in STPCL scores from striker through petty officer third class to petty officer second class.

Conclusions

The following conclusions seem justified by the results of the research described in this report:

1. Skills involved in the Naval ratings of aviation electrician's mate, aviation electronics technician, aviation fire control technician, and TRADEVMAN are scaled by both the Thurstone method of equal-appearing intervals and the Guttman method of scalogram analysis.

2. It is possible to construct a single scaled technical proficiency check list which can be applied to technicians in any of four electronics ratings.

3. As measured by the Scaled Technical Proficiency Check List, TRADEVMEN are significantly less proficient on the electronics tasks common to all four ratings than aviation electrician's mates, aviation electronics technicians, and aviation fire control technicians.

4. As reflected by the scaled lists, the proficiency of electronically oriented technicians on the tasks done in their ratings rises as they are promoted in pay grade from striker to petty officer second class.

5. The Naval attitudes of electronics technicians, as expressed in a self-report questionnaire, generally are not strongly related to technical fleet proficiency, although attitudes regarding certain aspects of the job may have slight, positive relationships with technical proficiency.

6. The measured fleet effectiveness of electronics technicians is not related to scores on the Navy Basic Test Battery or to technical school grades.

As a result of three Applied Psychological Services' studies, it seems reasonable to state that technical skills involved in Naval ratings are scalable in the same manner as attitudes and the sensory phenomena which have been previously scaled psychophysically and that the scaling can be established either within a single rating or across several related ratings. The scales appear to offer a quick, convenient way of evaluating the post-training technical proficiency of Naval personnel and to provide one basis for judging the effectiveness of technical training programs.

REFERENCES

Green, B. F. Attitude measurement. In G. Lindzey (Ed.), Handbook of social psychology. Cambridge, Mass.: Addison Wesley, 1954. Pp. 335-369.

Green, B. F. A method of scalogram analysis using summary statistics. Psychometrika, 1956, 21, 79-88.

Guttman, L. The basis for scalogram analysis. In S. A. Stouffer *et al.*, Measurement and prediction. Princeton, N. J.: Princeton University Press, 1950. Pp. 60-90.

Richlin, M., Federman, P., and Siegel, A. I. Post-training performance criterion development and application: development and application of a TBCL criterion to the SESR program for jet aviation machinist's mates. Wayne, Pa.: Applied Psychological Services, 1958.

Richlin, M., Siegel, A. I., and Schultz, D. G. Post-training performance criterion development and application: development and application of a TBCL criterion to the SESR program for aviation electronics technicians. Wayne, Pa.: Applied Psychological Services, 1960.

Siegel, A. I., Richlin, M., and Federman, P. Post-training performance criterion development and application: development and application of TBCL criteria to the SESR program for the air controlman and the parachute rigger ratings. Wayne, Pa.: Applied Psychological Services, 1958.

Siegel, A. I., and Benson, S. Post-training performance criterion development and application: technical performance check list criteria which meet the Thurstone and Guttman scalability requirements. Wayne, Pa.: Applied Psychological Services, 1959.

Siegel, A. I., Richlin, M., and Federman, P. A comparative study of "transfer through generalization" and "transfer through identical elements" in technical training. J. appl. Psychol., 1960, 44, 27-30.

Siegel, A. I., Schultz, D. G., and Benson, S. Post-training performance criterion development and application: a further study into technical performance check list criteria which meet the Thurstone and Guttman scalability requirements. Wayne, Pa.: Applied Psychological Services, 1960.

Siegel, A. I., and Schultz, D. G. Post-training performance criterion development and application: the measurement of motivation and its relationship to fleet proficiency. Wayne, Pa.: Applied Psychological Services, 1960.

Suchman, E. A. The utility of scalogram analysis. In S. A. Stouffer *et al.*, Measurement and prediction. Princeton, N. J.: Princeton University Press, 1950, Pp. 122-171.

Thorndike, R. L., and Hagen, Elizabeth. 10,000 careers. New York: John Wiley, 1959.

Thurstone, L. L., and Chave, E. J. The measurement of attitude. Chicago: University of Chicago Press, 1929.

Torgerson, W. S. Theory and methods of scaling. New York: John Wiley, 1958.

APPENDIX A

Appendix A presents the preliminary task list.

FORM 42

Your Name ____________________

Rate ____________________

Number of years in rate ____________

Squadron ____________________

Location ____________________

Prepared by

Applied Psychological Services

Wayne, Pa.

under

Contract Nonr-2279(00)

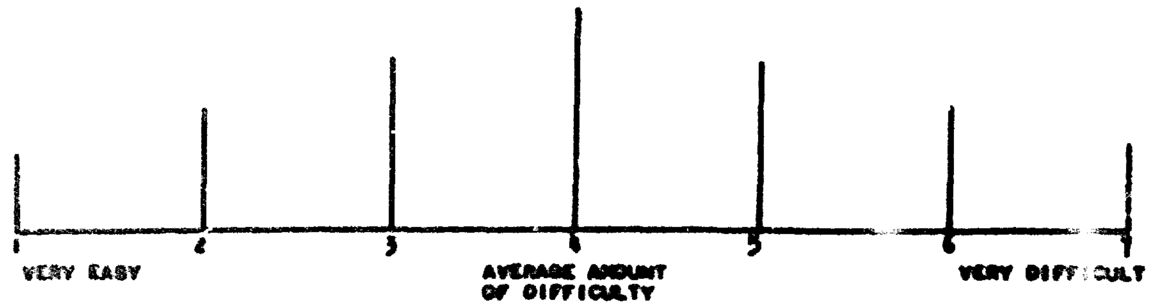
with the

Office of Naval Research

August 1959

DIRECTIONS

Listed below you will find 28 tasks that are done by strikers in your rating. Read over the 28 tasks and then using the scale below as a guide indicate the degree of difficulty you believe that the typical striker encounters on each of the tasks before he can perform it proficiently. These judgments are to be made in the following manner:



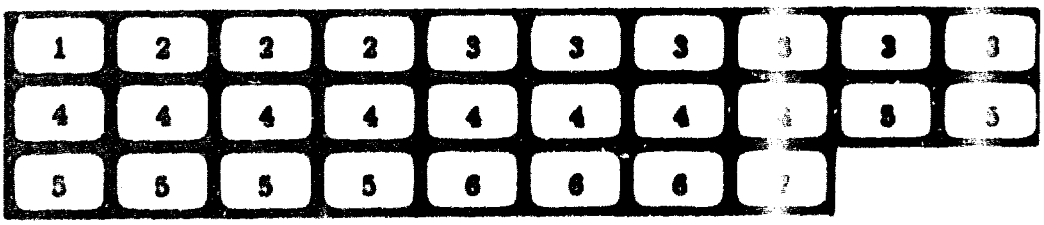
1. Using the stickers provided place the sticker with a 1 on it next to the job that is least difficult for the average striker.

2. Then pick the three jobs that fall in category "two" and put the three stickers with 2's on them next to these items.

3. Now find the job that is the most difficult and place the sticker with the 7 on it next to this item.

4. Find the three items you want to place in category "six" and put these stickers next to these three items.

5. Now with the tasks remaining place them in their proper category by putting the sticker (i.e., 3, 4, or 5) with the category number next to the items.



JOB ASSIGNMENTS	SCALE VALUE
1. Operating	
2. Preflight inspecting	
3. Postflight inspecting	
4. Inflight inspecting	
5. Periodically inspecting	
6. Maintaining	
7. Removing	
8. Repairing	
9. Replacing	
10. Performing preventative maintenance	
11. Trouble shooting/isolating malfunction(s) in	
12. Calibrating	
13. Aligning	
14. Following block diagrams for	
15. Using standard schematics for	
16. Analyzing standard circuitry in	
17. Employing safety precautions on	
18. Using proper safety precautions for self when working on	
19. Making out failure reports for	
20. Using manuals of a technical nature for	

JOB ASSIGNMENTS	SCALE VALUE
21. Using appropriate test equipment for determining malfunctions in the	
22. Using ASO catalogue for replacement parts for	
23. Using mathematical formulas necessary for solving circuit equations for	
24. Employing electronic principles involved in maintenance of	
25. Knowing relationship of equipment to other related	
26. Instructing others in operation of	
27. Instructing others in maintenance of	
28. Instructing others in the inspection of	

Check to make sure that you have answered every question.

APPENDIX B

Appendix B presents the evaluation form.

FORM 44

Name of man you are evaluating ______________________

Rate ______________________

Squadron ______________________

Location ______________________

Your Name and Rate ______________________

Number of months you have supervised man you are evaluating ______

SCALED TECHNICAL PROFICIENCY CHECK LIST

Prepared by

Applied Psychological Services

Wayne, Pa.

under

Contract Nonr-2279(00)

with the

Office of Naval Research

August 1959

DIRECTIONS

For each of the tasks listed indicate (by placing a check mark in the appropriate column) whether or not the man you are evaluating has been checked out as being proficient (i.e., is he capable of doing the task "on his own" without direct supervision?) If the man you are rating has not been checked out as being proficient on a task, because he has not worked on this task, this should be indicated by a check mark in the third column.

Examples:

1. If the man has been checked out on a task place a check mark (✓) in column one (1).
2. If he has not been checked out as being proficient but has actually worked on this task place a check mark (✓) in column two (2).
3. If he has not been checked out as being proficient because he has not performed this task, place a check mark (✓) in column three (3) next to the appropriate item.

Be sure that you give an answer for every task that is listed. You may not be certain about some items, but answer every item on the basis of what appears to be the most nearly correct answer for the man you are rating.

THE MAN BEING EVALUATED ON THE FOLLOWING ITEMS IS: 1. CHECKED OUT; 2. NOT CHECKED OUT; OR 3. HAS NOT WORKED ON MOST OF THE EQUIPMENT INVOLVED IN HIS RATING AND FOUND IN THIS SQUADRON.

	HAS WORKED ON TASK AND IS: CHECKED OUT 1	HAS WORKED ON TASK AND IS: NOT CHECKED OUT 2	HAS NOT WORKED ON TASK 3
1. Operating			
2. Preflight inspecting			
3. Postflight inspecting			
4. Inflight inspecting			
5. Periodically inspecting			
6. Maintaining			
7. Removing			
8. Repairing			
9. Replacing			
10. Performing preventative maintenance			
11. Trouble shooting/isolating malfunction(s) in			
12. Calibrating			
13. Aligning			
14. Following block diagrams for			
15. Using standard schematics for			
16. Analyzing standard circuitry in			
17. Employing safety precautions on			
18. Using proper safety precautions for self when working on			
19. Making out failure reports for			
20. Using manuals of a technical nature for			

	HAS WORKED ON TASK AND IS		HAS NOT WORKED ON TASK
	CHECKED OUT	NOT CHECKED OUT	
	1	2	3
21. Using appropriate test equipment for determining malfunctions in the			
22. Using ASO catalogue for replacement parts for			
23. Using mathematical formulas necessary for solving circuit equations for			
24. Employing electronic principles involved in maintenance of			
25. Knowing relationship of equipment to other related			
26. Instructing others in operation of			
27. Instructing others in maintenance of			
28. Instructing others in the inspection of			